



SEQUENCE LISTING

<110> Chiang, Vincent L.
Wu, Luguang
Joshi, Chandrashekhar P.
Carraway, Daniel T.

<120> METHOD FOR ENHANCING CELLULOSE AND MODIFYING LIGNIN BIOSYNTHESIS
IN PLANTS

<130> 066040-9675

<140> 09/980,043

<141> 2002-04-05

<150> PCT/US00/13637

<151> 2000-05-18

<150> 60/135,280

<151> 1999-05-21

<160> 7

<170> Word 97 (MS-DOS text format)

<210> 1

<211> 3232

<212> DNA

<213> Populus tremuloides

<220>

<221> CDS

<222> (69)..(3002)

<400> 1

gtcgaccac gcgtccgtct tgaagaata tgaagttgta aagagctggt aaagtggtaa 60

taagcaag atg atg gaa tct ggg gct cct ata tgc cat acc tgt ggt gaa 110
Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu
1 5 10

cag gtg ggg cat gat gca aat ggg gag cta ttt gtg gct tgc cat gag 158
Gln Val Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu
15 20 25 30

tgt agc tat ccc atg tgc aag tct tgt ttc gag ttt gaa atc aat gag 206
Cys Ser Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Asn Glu
35 40 45

ggc cgg aaa gtt tgc ttg cgg tgt ggc tcg cca tat gat gag aac ttg 254
Gly Arg Lys Val Cys Leu Arg Cys Gly Ser Pro Tyr Asp Glu Asn Leu
50 55 60

ctg gat gat gta gaa aag aag ggg tct ggc aat caa tcc aca atg gca 302
Leu Asp Asp Val Glu Lys Lys Gly Ser Gly Asn Gln Ser Thr Met Ala
65 70 75

tct cac ctc aac gat tct cag gat gtc gga atc cat gct aga cat atc 350
Ser His Leu Asn Asp Ser Gln Asp Val Gly Ile His Ala Arg His Ile
80 85 90

agt agt gtg tcc act gtg gat agt gaa atg aat gat gaa tat ggg aat	398
Ser Ser Val Ser Thr Val Asp Ser Glu Met Asn Asp Glu Tyr Gly Asn	
95 100 105 110	
cca att tgg aag aat cgg gtg aag agc tgt aag gat aaa gag aac aag	446
Pro Ile Trp Lys Asn Arg Val Lys Ser Cys Lys Asp Lys Glu Asn Lys	
115 120 125	
aag aaa aag aga agt cct aag gct gaa act gaa cca gct caa gtt cct	494
Lys Lys Lys Arg Ser Pro Lys Ala Glu Thr Glu Pro Ala Gln Val Pro	
130 135 140	
aca gaa cag cag atg gaa gag aaa ccg tct gca gag gct tcg gag ccg	542
Thr Glu Gln Gln Met Glu Glu Lys Pro Ser Ala Glu Ala Ser Glu Pro	
145 150 155	
ctt tca att gtt tat cca att cca cgc aac aag ctc aca cca tac aga	590
Leu Ser Ile Val Tyr Pro Ile Pro Arg Asn Lys Leu Thr Pro Tyr Arg	
160 165 170	
gca gtg atc att atg cga ctg gtc att ctg ggc ctc ttc ttc cac ttc	638
Ala Val Ile Ile Met Arg Leu Val Ile Leu Gly Leu Phe Phe His Phe	
175 180 185 190	
aga ata aca aat cct gtc gat agt gcc ttt ggc ctg tgg ctt act tct	686
Arg Ile Thr Asn Pro Val Asp Ser Ala Phe Gly Leu Trp Leu Thr Ser	
195 200 205	
gtc ata tgt gag atc tgg ttt gca ttt tct tgg gtg ttg gat cag ttc	734
Val Ile Cys Glu Ile Trp Phe Ala Phe Ser Trp Val Leu Asp Gln Phe	
210 215 220	
ccc aag tgg aat cct gtc aat aga gaa acg tat atc gaa agg ctg tcg	782
Pro Lys Trp Asn Pro Val Asn Arg Glu Thr Tyr Ile Glu Arg Leu Ser	
225 230 235	
gca agg tat gaa aga gag ggt gag cct tct cag ctt gct ggt gtg gat	830
Ala Arg Tyr Glu Arg Glu Gly Glu Pro Ser Gln Leu Ala Gly Val Asp	
240 245 250	
ttt ttc gtg agt act gtt gat ccg ctg aag gaa ccg cca ttg atc act	878
Phe Phe Val Ser Thr Val Asp Pro Leu Lys Glu Pro Pro Leu Ile Thr	
255 260 265 270	
gcc aat aca gtc ctt tcc atc ctt gct gtg gac tat ccc gtc gat aaa	926
Ala Asn Thr Val Leu Ser Ile Leu Ala Val Asp Tyr Pro Val Asp Lys	
275 280 285	
gtc tcc tgc tac gtg tct gat gat ggt gca gct atg ctt tca ttt gaa	974
Val Ser Cys Tyr Val Ser Asp Asp Gly Ala Ala Met Leu Ser Phe Glu	
290 295 300	
tct ctt gta gaa aca gct gag ttt gca agg aag tgg gtt ccg ttc tgc	1022
Ser Leu Val Glu Thr Ala Glu Phe Ala Arg Lys Trp Val Pro Phe Cys	
305 310 315	
aaa aaa ttc tca att gaa cca aga gca ccg gag ttt tac ttc tca cag	1070
Lys Lys Phe Ser Ile Glu Pro Arg Ala Pro Glu Phe Tyr Phe Ser Gln	
320 325 330	

aaa att gat tac ttg aaa gac aag gtt caa cct tct ttc gtg aaa gaa	1118
Lys Ile Asp Tyr Leu Lys Asp Lys Val Gln Pro Ser Phe Val Lys Glu	
335 340 345 350	
cg t aga gca atg aaa agg gat tat gaa gag tac aaa gtc cga gtt aat	1166
Arg Arg Ala Met Lys Arg Asp Tyr Glu Glu Tyr Lys Val Arg Val Asn	
355 360 365	
gcc ctg gta gca aag gct cag aaa aca cct gaa gaa gga tgg act atg	1214
Ala Leu Val Ala Lys Ala Gln Lys Thr Pro Glu Glu Gly Trp Thr Met	
370 375 380	
caa gat gga aca cct tgg cct ggg aat aac aca cgt gat cac cct ggg	1262
Gln Asp Gly Thr Pro Trp Pro Gly Asn Asn Thr Arg Asp His Pro Gly	
385 390 395	
cat gat tca ggt ctt cct tgg gaa ata ctg gga gct cgt gac att gaa	1310
His Asp Ser Gly Leu Pro Trp Glu Ile Leu Gly Ala Arg Asp Ile Glu	
400 405 410	
gga aat gaa cta cct cgt cta gta tat gtc tcc agg gag aag aga cct	1358
Gly Asn Glu Leu Pro Arg Leu Val Tyr Val Ser Arg Glu Lys Arg Pro	
415 420 425 430	
ggc tac cag cac cac aaa aag gct ggt gca gaa aat gct ctg gtg aga	1406
Gly Tyr Gln His His Lys Lys Ala Gly Ala Glu Asn Ala Leu Val Arg	
435 440 445	
gtg tct gca gta ctc aca aat gct ccc tac atc ctc aat gtt gat tgt	1454
Val Ser Ala Val Leu Thr Asn Ala Pro Tyr Ile Leu Asn Val Asp Cys	
450 455 460	
gat cac tat gta aac aat agc aag gct gtt cga gag gca atg tgc atc	1502
Asp His Tyr Val Asn Asn Ser Lys Ala Val Arg Glu Ala Met Cys Ile	
465 470 475	
ctg atg gac cca caa gta ggt cga gat gta tgc tat gtg cag ttc cct	1550
Leu Met Asp Pro Gln Val Gly Arg Asp Val Cys Tyr Val Gln Phe Pro	
480 485 490	
cag agg ttt gat ggc ata gat aag agt gat cgc tac gcc aat cgt aac	1598
Gln Arg Phe Asp Gly Ile Asp Lys Ser Asp Arg Tyr Ala Asn Arg Asn	
495 500 505 510	
gta gtt ttc ttt gat gtt aac atg aaa ggg ttg gat ggc att caa gga	1646
Val Val Phe Phe Asp Val Asn Met Lys Gly Leu Asp Gly Ile Gln Gly	
515 520 525	
cca gta tac gta gga act ggt tgt gtt ttc aac agg caa gca ctt tac	1694
Pro Val Tyr Val Gly Thr Gly Cys Val Phe Asn Arg Gln Ala Leu Tyr	
530 535 540	
ggc tac ggg cct cct tct atg ccc agc tta cgc aag aga aag gat tct	1742
Gly Tyr Gly Pro Pro Ser Met Pro Ser Leu Arg Lys Arg Lys Asp Ser	
545 550 555	
tca tcc tgc ttc tca tgt tgc tgc ccc tca aag aag aag cct gct caa	1790
Ser Ser Cys Phe Ser Cys Cys Cys Pro Ser Lys Lys Lys Pro Ala Gln	
560 565 570	
gat cca gct gag gta tac aga gat gca aaa aga gag gat ctc aat gct	1838

Asp 575	Pro	Ala	Glu	Val	Tyr 580	Arg	Asp	Ala	Lys	Arg 585	Glu	Asp	Leu	Asn	Ala 590		
gcc	ata	ttt	aat	ctt	aca	gag	att	gat	aat	tat	gac	gag	cat	gaa	agg	1886	
Ala	Ile	Phe	Asn	Leu	Thr	Glu	Ile	Asp	Asn	Tyr	Asp	Glu	His	Glu	Arg		
				595					600					605			
tca	atg	ctg	atc	tcc	cag	ttg	agc	ttt	gag	aaa	act	ttt	ggc	tta	tct	1934	
Ser	Met	Leu	Ile	Ser	Gln	Leu	Ser	Phe	Glu	Lys	Thr	Phe	Gly	Leu	Ser		
			610					615					620				
tct	gtc	ttc	att	gag	tct	aca	cta	atg	gag	aat	gga	gga	gta	ccc	gag	1982	
Ser	Val	Phe	Ile	Glu	Ser	Thr	Leu	Met	Glu	Asn	Gly	Gly	Val	Pro	Glu		
		625					630					635					
tct	gcc	aac	tca	cca	cca	ttc	atc	aag	gaa	gcg	att	caa	gtc	atc	ggc	2030	
Ser	Ala	Asn	Ser	Pro	Pro	Phe	Ile	Lys	Glu	Ala	Ile	Gln	Val	Ile	Gly		
	640					645					650						
tgt	ggc	tat	gaa	gag	aag	act	gaa	tgg	gga	aaa	cag	att	ggg	tgg	ata	2078	
Cys	Gly	Tyr	Glu	Glu	Lys	Thr	Glu	Trp	Gly	Lys	Gln	Ile	Gly	Trp	Ile		
655					660					665					670		
tat	ggg	tca	gtc	act	gag	gat	atc	tta	agt	ggc	ttc	aag	atg	cac	tgc	2126	
Tyr	Gly	Ser	Val	Thr	Glu	Asp	Ile	Leu	Ser	Gly	Phe	Lys	Met	His	Cys		
				675					680					685			
cga	gga	tgg	aga	tca	att	tac	tgc	atg	ccc	gta	agg	cct	gca	ttc	aaa	2174	
Arg	Gly	Trp	Arg	Ser	Ile	Tyr	Cys	Met	Pro	Val	Arg	Pro	Ala	Phe	Lys		
			690					695					700				
gga	tct	gca	ccc	atc	aac	ctg	tct	gat	aga	ttg	cac	cag	gtc	ctc	cga	2222	
Gly	Ser	Ala	Pro	Ile	Asn	Leu	Ser	Asp	Arg	Leu	His	Gln	Val	Leu	Arg		
		705					710					715					
tgg	gct	ctt	ggg	tct	gtg	gaa	att	ttc	ttt	agc	aga	cac	tgt	ccc	ctc	2270	
Trp	Ala	Leu	Gly	Ser	Val	Glu	Ile	Phe	Phe	Ser	Arg	His	Cys	Pro	Leu		
	720					725					730						
tgg	tac	ggg	ttt	gga	gga	ggc	cgt	ctt	aaa	tgg	ctc	caa	agg	ctt	gcg	2318	
Trp	Tyr	Gly	Phe	Gly	Gly	Gly	Arg	Leu	Lys	Trp	Leu	Gln	Arg	Leu	Ala		
735				740					745						750		
tat	ata	aac	acc	att	gtg	tac	cca	ttt	aca	tcc	ctc	cct	ctc	att	gcc	2366	
Tyr	Ile	Asn	Thr	Ile	Val	Tyr	Pro	Phe	Thr	Ser	Leu	Pro	Leu	Ile	Ala		
				755					760					765			
tat	tgc	aca	att	cct	gca	gtt	tgt	ctg	ctc	acc	gga	aaa	ttc	atc	ata	2414	
Tyr	Cys	Thr	Ile	Pro	Ala	Val	Cys	Leu	Leu	Thr	Gly	Lys	Phe	Ile	Ile		
			770					775					780				
cca	acg	ctc	tca	aac	ctg	gca	agc	atg	ctg	ttt	ctt	ggc	ctc	ttt	atc	2462	
Pro	Thr	Leu	Ser	Asn	Leu	Ala	Ser	Met	Leu	Phe	Leu	Gly	Leu	Phe	Ile		
		785					790					795					
tcc	atc	att	gta	act	gcg	gtg	ctt	gag	cta	aga	tgg	agc	ggg	gtc	agc	2510	
Ser	Ile	Ile	Val	Thr	Ala	Val	Leu	Glu	Leu	Arg	Trp	Ser	Gly	Val	Ser		
	800					805					810						
att	gaa	gat	tta	tgg	cgt	aat	gaa	caa	ttc	tgg	gtg	atc	gga	ggg	gtt	2558	
Ile	Glu	Asp	Leu	Trp	Arg	Asn	Glu	Gln	Phe	Trp	Val	Ile	Gly	Gly	Val		

815	820	825	830	
tca gcc cat ctc ttt gcg gtc ttc cag gga ttc tta aaa atg ttg gct				2606
Ser Ala His Leu Phe	Ala Val Phe	Gln Gly Phe Leu Lys	Met Leu Ala	
	835	840	845	
ggc atc gat acg aac ttc act gtc aca gca aaa gca gcc gaa gat gca				2654
Gly Ile Asp Thr Asn Phe Thr Val Thr Ala Lys Ala Ala Glu Asp Ala				
	850	855	860	
gaa ttt ggg gag cta tat atg gtc aag tgg aca aca ctt ttg att cct				2702
Glu Phe Gly Glu Leu Tyr Met Val Lys Trp Thr Thr Leu Leu Ile Pro				
	865	870	875	
cca acc aca ctt ctc att atc aat atg tcg ggt tgt gct gga ttc tct				2750
Pro Thr Thr Leu Leu Ile Ile Asn Met Ser Gly Cys Ala Gly Phe Ser				
	880	885	890	
gat gca ctc aac aaa gga tat gaa gca tgg ggg cct ctc ttt ggc aag				2798
Asp Ala Leu Asn Lys Gly Tyr Glu Ala Trp Gly Pro Leu Phe Gly Lys				
	895	900	905	910
gtg ttc ttt gct ttc tgg gtg att ctt cat ctc tat cca ttc ctt aaa				2846
Val Phe Phe Ala Phe Trp Val Ile Leu His Leu Tyr Pro Phe Leu Lys				
	915	920	925	
ggt cta atg ggt cgc caa aac cta aca cca acc att gtt gtt ctc tgg				2894
Gly Leu Met Gly Arg Gln Asn Leu Thr Pro Thr Ile Val Val Leu Trp				
	930	935	940	
tca gtg ctg ttg gcc tct gtc ttc tct ctc gtt tgg gtc aag atc aat				2942
Ser Val Leu Leu Ala Ser Val Phe Ser Leu Val Trp Val Lys Ile Asn				
	945	950	955	
cca ttc gtt aac aaa gtt gat aac acc ttg gtt gcg gag acc tgc att				2990
Pro Phe Val Asn Lys Val Asp Asn Thr Leu Val Ala Glu Thr Cys Ile				
	960	965	970	
tcc att gat tgc tgagctacct ccaataagtc tctcccagta ttttgggggtt				3042
Ser Ile Asp Cys				
	975			
acaaaacctt tgggaattgg aatatgatcc tcgtttagt ttccctcaag aaagcacata				3102
tcgctgtcag tattttaaag aactgcaaga tgattgttct ctatgaagtt ttgaacagtt				3162
tgaaatgata ttatgttaaa atacagggtt tgattgtgtt gaaaaaaaaa aagaaaaaaaaa				3222
aaaaaaaaaa				3232

<210> 2

<211> 978

<212> PRT

<213> Populus tremuloides

<400> 2

Met	Met	Glu	Ser	Gly	Ala	Pro	Ile	Cys	His	Thr	Cys	Gly	Glu	Gln	Val
1				5				10					15		

Gly	His	Asp	Ala 20	Asn	Gly	Glu	Leu	Phe 25	Val	Ala	Cys	His	Glu 30	Cys	Ser
Tyr	Pro	Met 35	Cys	Lys	Ser	Cys	Phe 40	Glu	Phe	Glu	Ile	Asn 45	Glu	Gly	Arg
Lys	Val 50	Cys	Leu	Arg	Cys	Gly 55	Ser	Pro	Tyr	Asp	Glu 60	Asn	Leu	Leu	Asp
Asp 65	Val	Glu	Lys	Lys	Gly 70	Ser	Gly	Asn	Gln	Ser 75	Thr	Met	Ala	Ser	His 80
Leu	Asn	Asp	Ser	Gln 85	Asp	Val	Gly	Ile	His 90	Ala	Arg	His	Ile	Ser 95	Ser
Val	Ser	Thr 100	Val	Asp	Ser	Glu	Met	Asn 105	Asp	Glu	Tyr	Gly	Asn 110	Pro	Ile
Trp	Lys	Asn 115	Arg	Val	Lys	Ser	Cys 120	Lys	Asp	Lys	Glu	Asn 125	Lys	Lys	Lys
Lys	Arg 130	Ser	Pro	Lys	Ala	Glu 135	Thr	Glu	Pro	Ala	Gln 140	Val	Pro	Thr	Glu
Gln 145	Gln	Met	Glu	Glu	Lys 150	Pro	Ser	Ala	Glu	Ala 155	Ser	Glu	Pro	Leu	Ser 160
Ile	Val	Tyr	Pro	Ile 165	Pro	Arg	Asn	Lys	Leu 170	Thr	Pro	Tyr	Arg	Ala 175	Val
Ile	Ile	Met 180	Arg	Leu	Val	Ile	Leu	Gly 185	Leu	Phe	Phe	His	Phe 190	Arg	Ile
Thr	Asn 195	Pro	Val	Asp	Ser	Ala	Phe 200	Gly	Leu	Trp	Leu	Thr 205	Ser	Val	Ile
Cys 210	Glu	Ile	Trp	Phe	Ala 215	Phe	Ser	Trp	Val	Leu	Asp 220	Gln	Phe	Pro	Lys
Trp 225	Asn	Pro	Val	Asn	Arg 230	Glu	Thr	Tyr	Ile	Glu 235	Arg	Leu	Ser	Ala	Arg 240
Tyr	Glu	Arg	Glu	Gly 245	Glu	Pro	Ser	Gln	Leu 250	Ala	Gly	Val	Asp 255	Phe	Phe
Val	Ser	Thr 260	Val	Asp	Pro	Leu	Lys	Glu 265	Pro	Pro	Leu	Ile	Thr 270	Ala	Asn
Thr	Val 275	Leu	Ser	Ile	Leu	Ala	Val 280	Asp	Tyr	Pro	Val	Asp 285	Lys	Val	Ser
Cys 290	Tyr	Val	Ser	Asp	Asp	Gly 295	Ala	Ala	Met	Leu	Ser 300	Phe	Glu	Ser	Leu
Val 305	Glu	Thr	Ala	Glu	Phe 310	Ala	Arg	Lys	Trp	Val 315	Pro	Phe	Cys	Lys	Lys 320
Phe	Ser	Ile	Glu	Pro 325	Arg	Ala	Pro	Glu	Phe 330	Tyr	Phe	Ser	Gln 335	Lys	Ile
Asp	Tyr	Leu	Lys	Asp	Lys	Val	Gln	Pro	Ser	Phe	Val	Lys	Glu	Arg	Arg

340					345					350					
Ala	Met	Lys	Arg	Asp	Tyr	Glu	Glu	Tyr	Lys	Val	Arg	Val	Asn	Ala	Leu
		355					360					365			
Val	Ala	Lys	Ala	Gln	Lys	Thr	Pro	Glu	Glu	Gly	Trp	Thr	Met	Gln	Asp
		370					375					380			
Gly	Thr	Pro	Trp	Pro	Gly	Asn	Asn	Thr	Arg	Asp	His	Pro	Gly	His	Asp
							390					395			400
Ser	Gly	Leu	Pro	Trp	Glu	Ile	Leu	Gly	Ala	Arg	Asp	Ile	Glu	Gly	Asn
				405					410					415	
Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	Gly	Tyr
			420					425					430		
Gln	His	His	Lys	Lys	Ala	Gly	Ala	Glu	Asn	Ala	Leu	Val	Arg	Val	Ser
			435				440					445			
Ala	Val	Leu	Thr	Asn	Ala	Pro	Tyr	Ile	Leu	Asn	Val	Asp	Cys	Asp	His
		450					455					460			
Tyr	Val	Asn	Asn	Ser	Lys	Ala	Val	Arg	Glu	Ala	Met	Cys	Ile	Leu	Met
				465			470					475			480
Asp	Pro	Gln	Val	Gly	Arg	Asp	Val	Cys	Tyr	Val	Gln	Phe	Pro	Gln	Arg
				485					490					495	
Phe	Asp	Gly	Ile	Asp	Lys	Ser	Asp	Arg	Tyr	Ala	Asn	Arg	Asn	Val	Val
			500					505					510		
Phe	Phe	Asp	Val	Asn	Met	Lys	Gly	Leu	Asp	Gly	Ile	Gln	Gly	Pro	Val
		515					520					525			
Tyr	Val	Gly	Thr	Gly	Cys	Val	Phe	Asn	Arg	Gln	Ala	Leu	Tyr	Gly	Tyr
		530					535					540			
Gly	Pro	Pro	Ser	Met	Pro	Ser	Leu	Arg	Lys	Arg	Lys	Asp	Ser	Ser	Ser
				545			550					555			560
Cys	Phe	Ser	Cys	Cys	Cys	Pro	Ser	Lys	Lys	Lys	Pro	Ala	Gln	Asp	Pro
				565					570					575	
Ala	Glu	Val	Tyr	Arg	Asp	Ala	Lys	Arg	Glu	Asp	Leu	Asn	Ala	Ala	Ile
			580					585					590		
Phe	Asn	Leu	Thr	Glu	Ile	Asp	Asn	Tyr	Asp	Glu	His	Glu	Arg	Ser	Met
		595					600					605			
Leu	Ile	Ser	Gln	Leu	Ser	Phe	Glu	Lys	Thr	Phe	Gly	Leu	Ser	Ser	Val
		610					615					620			
Phe	Ile	Glu	Ser	Thr	Leu	Met	Glu	Asn	Gly	Gly	Val	Pro	Glu	Ser	Ala
				625			630					635			640
Asn	Ser	Pro	Pro	Phe	Ile	Lys	Glu	Ala	Ile	Gln	Val	Ile	Gly	Cys	Gly
				645					650					655	
Tyr	Glu	Glu	Lys	Thr	Glu	Trp	Gly	Lys	Gln	Ile	Gly	Trp	Ile	Tyr	Gly
			660					665					670		

[illegible]

<210> 3
 <211> 1010
 <212> DNA
 <213> Populus tremuloides

<220>
 <221> CDS
 <222> (841)..(1008)

<220>
 <221> misc_feature
 <223> 5' flanking region of PtCelA coding sequence

<400> 3

```

gaattcgccc ttttgaattc aggagacgat agtttcgggt tcgttgaatg gctttgttca 60
cttctgggtct agcaatttgc aaaagaagtt acaaaacaaa tgcatattat gtaaatttaa 120
caagagatgg gttctatggg cacttattta tgcccatcat ttgttctggg gttactcttt 180
atagtctgat tcgaagttgc aaactgccgt ttctgggtatt gcaattatgt agccataaac 240
tgtaaatcct gtagctatta gcggaccaac aaccagatat acgggatcag cgtcgtaaaa 300
gagatctcca ttctacgttt ctttctaatt tttccgtttc agtgagagaa ttaccctgat 360
acattgacat gatgattgat gattatggga accattccga tgttagacac gagaccatct 420
ggatcctgcc agttttcagt tcacatggca tctcagccca agatcatgtg tttatacgcc 480
taatgacttg tattgaaagt ttggtaagtt gaagatgtgc tctgccaac agaaaccttc 540
cttaaatttc cagcaaatct ttcaaacttg gccttacacc ccgaaaatag acgtgcttct 600
acttgggttc ttggaaacca tgcaccaacc gccatacccc accaaccac caccctcaac 660
cttctcttcg ccattacaaa aatgtcagta ccaccctctg aaagacacca acacacccta 720
gctttggtta gggatatttg tataaaaaca aggccaaaac aaaagattgg aaggaagcag 780
aggaagaccc tcttgaaaga attgaagttg taaagagctg gtaaagtggg aataagcaag 840

atg atg gaa tct ggg gct cct ata tgc cat acc tgt ggt gaa cag gtg 888
Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu Gln Val
  1             5             10             15

ggg cat gat gca aat ggg gag cta ttt gtg gct tgc cat gag tgt agc 936
Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu Cys Ser
          20             25             30

tat ccc atg tgc aag tct tgt ttc gag ttt gaa atc aaa gag ggc cgg 984
Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Lys Glu Gly Arg
          35             40             45

aaa gtt tgc ttg cgg tgt ggc tcg ag 1010
Lys Val Cys Leu Arg Cys Gly Ser
  50             55

```

<210> 4

<211> 3444

<212> DNA

<213> *Arabidopsis thaliana*

<400> 4

```
gcggccgcgg ttaatcgccg gttctcacia caggaatgag tttgtcctca ttaatgccga 60
tgagaatgcc cgaataagat cagtccaaga gctgagtggg cagacatgtc aaatctgcag 120
agatgagatc gaattgactg ttgatggaga accgtttgtg gcatgtaacg aatgtgcatt 180
ccctgtgtgt agaccttgct atgagtacga aagacgagaa ggcaatcaag cttgtccaca 240
gtgcaaaacc cgtttcaaac gtcttaaagg aagtccaaga gttgaagggt atgaagagga 300
agatgacatt gatgatttag acaatgagtt tgagtatgga aataatggga ttggatttga 360
tcaggtttct gaaggatagt caatctctcg tcgcaactcc ggtttccac aatctgattt 420
ggattcagct ccacctggct ctgagattcc attgctgact tacggcgacg aggacgttga 480
gatttcttct gatagacatg ctcttattgt tcctccttca cttgggtggc atggcaatag 540
agttcatcct gtttctcttt ctgacccgac cgtggctgca catcgaaggc tgatggtacc 600
tcagaaagat cttgcggttt atggttatgg aagtgtcgct tggaaagatc ggatggagga 660
atggaagaga aagcagaatg agaaacttca ggttgtagg catgaaggag atcctgattt 720
tgaagatggt gatgatgtg attttccaat gatggatgag ggaaggcagc cattgtctat 780
gaagatacca atcaaatcga gcaagataaa tccttaccgg atgttaattg tgctacgtct 840
tgtgattctt ggtctcttct ttactaccg tattcttcac cccgtcaaag atgcatatgc 900
tttgtggctt atttctgtta tatgtgagat atggtttgct gtttcatggg ttcttgatca 960
gttcctaaa tggtacccta tcgagcgaga aacgtacttg gaccgactct cattaagata 1020
tgagaaagaa gggaaaccgt cgggactatc ccctgtggat gtatttgta gtacagtga 1080
tccattgaaa gagcctccgc ttattactgc aaatactgtc ttgtctatc ttgtgttga 1140
ttatcctgtc gataagggtt cttgttacgt atctgatgat ggtgctgcta tgcttacttt 1200
cgaagctctt tctgagaccg ctgaattcgc aaggaaatgg gttcctttct gcaagaaata 1260
ttgtatttag cctcgtgctc ccgaatggtt ttctgccat aaaatggact acttgaagaa 1320
taaagttcat cccgctattg ttagggagcg gcgagccatg aagagagatt atgaagaatt 1380
caaagtaaa atcaatgctt tagtagcaac agcacagaaa gtgcctgagg atggttggac 1440
tatgcaagac ggtacacctt ggcccggtta tagtgtgca gatcatcctg gcatgattca 1500
ggtcttctt ggaagtgcg gtgttcgtga tgcgaaaac aacgagttgc ctgattagt 1560
ttacgtttct cgtgagaaga gaccggatt tgatcaccat aagaaggctg gagctatgaa 1620
ttccctgata cgagtctctg gggttctatc aaatgctcct taccttctga atgtcgattg 1680
tgatcactac atcaacaata gcaaagctct tagagaagca atgtgtttca tgatggatcc 1740
tcagtcagga aagaaaatct gttatgttca gttccctcaa aggttcgatg ggattgatag 1800
gcagatcga tactcaaatc gcaatgttgt gttctttgat atcaatatga aaggtttgga 1860
tgggtacaa gggcctatat acgtcggtag aggttggtt ttcaggaggc aagcgcttta 1920
cggatttgat gcaccgaaga agaagaaggg ccacgtaag acatgcaatt gctggccaaa 1980
atggtgtctc ctatgttttg gttcaagaaa gaatcgtaaa gcaaagacag tggctgcgga 2040
taagaagaag aagaataggg aagcgtcaaa gcagatccac gcattagaaa atatcgaaga 2100
gggccgcggt cataaagttc ttaacgtaga acagtcaacc gaggcaatgc aaatgaagtt 2160
gcagaagaaa tatgggcagt ctctgtatt tgttgcatt gcgctctgg agaatggtgg 2220
gatggctaga aacgcaagcc cgcttgtct gcttaaagaa gccatccaag tcattagtcg 2280
cggatatgaa gataaaactg aatggggaaa agagattggg tggatctatg gttctgttac 2340
cgaagatatt cttacgggtt ctaagatgca ttctcatggt tggagacatg tttattgtac 2400
accaaagtta gcggctttca aaggatcagc tccaatcaat ctttcggatc gtctccatca 2460
agttcttcta tgggcgcttg ggtcgggtga gattttcttg agtaggcatt gtctatattg 2520
gtatggttat ggaggtgggt tgaaatggct tgagcggttg tcctacatta actctgtgg 2580
ttaccgctgg acctctctac cgctcatcgt ttactgttct ctccctgcca tctgtcttct 2640
cactgaaaaa ttcatcgttc ccgagattag caactatgcg agtatcctct tcatggcgct 2700
cttctcgtcg attgcaataa cgggtattct cgagatgcaa tggggcaaaag ttgggatcga 2760
tgattggtgg agaaacgaac agttttgggt cattggaggt gtttctgcgc atctgtttgc 2820
tctcttccaa ggtctcttca aggttcttgc tgggtgcgac actaacttca cagtcacatc 2880
aaaagcagct agttctgag agttctctga cctttacctc ttcaaatgga ctctacttct 2940
catccctcca atgactctac tcatcataaa cgtcattgga gtcatagtcg gactctctga 3000
tgccatcagc aatggatacg actcgtgggg accgcttttc ggaagactgt tctttgcact 3060
ttgggtcacc attcatcttt acccgcttct taaaggtttg cttgggaaac aagatagaat 3120
gccaaccatt attgtcgtct ggtccatcct cctggcctcg attcttacac ttctttgggt 3180
ccgggttaat ccgtttgtgg cgaaaggcgg tcctattctc gagatctgtg gtttagactg 3240
cttgtgattc gattgaccgg tggatgggtt ggtgaaaaag gtttaattcc cacggatcaa 3300
```

agagaggtaa gagagatatt gttttacctc taaaagactc cttcattgtg ttcattagat 3360
 gatgaaaaat gaaaagaaaa agaagattta attttggtac gagaattgtt atttttgcaa 3420
 gaatgtgttg tagatagcgg ccgc 3444

<210> 5
 <211> 1080
 <212> PRT
 <213> Arabidopsis thaliana

<220>
 <221> misc_feature
 <223> cellulose synthase

<400> 5

Arg	Pro	Arg	Leu	Ile	Ala	Gly	Ser	His	Asn	Arg	Asn	Glu	Phe	Val	Leu	1	5	10	15
Ile	Asn	Ala	Asp	Glu	Asn	Ala	Arg	Ile	Arg	Ser	Val	Gln	Glu	Leu	Ser	20	25	30	
Gly	Gln	Thr	Cys	Gln	Ile	Cys	Arg	Asp	Glu	Ile	Glu	Leu	Thr	Val	Asp	35	40	45	
Gly	Glu	Pro	Phe	Val	Ala	Cys	Asn	Glu	Cys	Ala	Phe	Pro	Val	Cys	Arg	50	55	60	
Pro	Cys	Tyr	Glu	Tyr	Glu	Arg	Arg	Glu	Gly	Asn	Gln	Ala	Cys	Pro	Gln	65	70	75	80
Cys	Lys	Thr	Arg	Phe	Lys	Arg	Leu	Lys	Gly	Ser	Pro	Arg	Val	Glu	Gly	85	90	95	
Asp	Glu	Glu	Glu	Asp	Asp	Ile	Asp	Asp	Leu	Asp	Asn	Glu	Phe	Glu	Tyr	100	105	110	
Gly	Asn	Asn	Gly	Ile	Gly	Phe	Asp	Gln	Val	Ser	Glu	Gly	Met	Ser	Ile	115	120	125	
Ser	Arg	Arg	Asn	Ser	Gly	Phe	Pro	Gln	Ser	Asp	Leu	Asp	Ser	Ala	Pro	130	135	140	
Pro	Gly	Ser	Gln	Ile	Pro	Leu	Leu	Thr	Tyr	Gly	Asp	Glu	Asp	Val	Glu	145	150	155	160
Ile	Ser	Ser	Asp	Arg	His	Ala	Leu	Ile	Val	Pro	Pro	Ser	Leu	Gly	Gly	165	170	175	
His	Gly	Asn	Arg	Val	His	Pro	Val	Ser	Leu	Ser	Asp	Pro	Thr	Val	Ala	180	185	190	
Ala	His	Arg	Arg	Leu	Met	Val	Pro	Gln	Lys	Asp	Leu	Ala	Val	Tyr	Gly	195	200	205	
Tyr	Gly	Ser	Val	Ala	Trp	Lys	Asp	Arg	Met	Glu	Glu	Trp	Lys	Arg	Lys	210	215	220	
Gln	Asn	Glu	Lys	Leu	Gln	Val	Val	Arg	His	Glu	Gly	Asp	Pro	Asp	Phe	225	230	235	240

Glu	Asp	Gly	Asp	Asp	Ala	Asp	Phe	Pro	Met	Met	Asp	Glu	Gly	Arg	Gln	
				245					250					255		
Pro	Leu	Ser	Met	Lys	Ile	Pro	Ile	Lys	Ser	Ser	Lys	Ile	Asn	Pro	Tyr	
			260					265					270			
Arg	Met	Leu	Ile	Val	Leu	Arg	Leu	Val	Ile	Leu	Gly	Leu	Phe	Phe	His	
		275					280					285				
Tyr	Arg	Ile	Leu	His	Pro	Val	Lys	Asp	Ala	Tyr	Ala	Leu	Trp	Leu	Ile	
	290					295					300					
Ser	Val	Ile	Cys	Glu	Ile	Trp	Phe	Ala	Val	Ser	Trp	Val	Leu	Asp	Gln	
305					310					315					320	
Phe	Pro	Lys	Trp	Tyr	Pro	Ile	Glu	Arg	Glu	Thr	Tyr	Leu	Asp	Arg	Leu	
				325					330					335		
Ser	Leu	Arg	Tyr	Glu	Lys	Glu	Gly	Lys	Pro	Ser	Gly	Leu	Ser	Pro	Val	
			340					345					350			
Asp	Val	Phe	Val	Ser	Thr	Val	Asp	Pro	Leu	Lys	Glu	Pro	Pro	Leu	Ile	
		355					360					365				
Thr	Ala	Asn	Thr	Val	Leu	Ser	Ile	Leu	Ala	Val	Asp	Tyr	Pro	Val	Asp	
		370				375					380					
Lys	Val	Ala	Cys	Tyr	Val	Ser	Asp	Asp	Gly	Ala	Ala	Met	Leu	Thr	Phe	
385					390					395					400	
Glu	Ala	Leu	Ser	Glu	Thr	Ala	Glu	Phe	Ala	Arg	Lys	Trp	Val	Pro	Phe	
				405					410					415		
Cys	Lys	Lys	Tyr	Cys	Ile	Glu	Pro	Arg	Ala	Pro	Glu	Trp	Tyr	Phe	Cys	
			420					425					430			
His	Lys	Met	Asp	Tyr	Leu	Lys	Asn	Lys	Val	His	Pro	Ala	Phe	Val	Arg	
		435					440					445				
Glu	Arg	Arg	Ala	Met	Lys	Arg	Asp	Tyr	Glu	Glu	Phe	Lys	Val	Lys	Ile	
	450					455					460					
Asn	Ala	Leu	Val	Ala	Thr	Ala	Gln	Lys	Val	Pro	Glu	Asp	Gly	Trp	Thr	
465					470					475					480	
Met	Gln	Asp	Gly	Thr	Pro	Trp	Pro	Gly	Asn	Ser	Val	Arg	Asp	His	Pro	
				485					490					495		
Gly	Met	Ile	Gln	Val	Phe	Leu	Gly	Ser	Asp	Gly	Val	Arg	Asp	Val	Glu	
			500					505					510			
Asn	Asn	Glu	Leu	Pro	Arg	Leu	Val	Tyr	Val	Ser	Arg	Glu	Lys	Arg	Pro	
		515					520					525				
Gly	Phe	Asp	His	His	Lys	Lys	Ala	Gly	Ala	Met	Asn	Ser	Leu	Ile	Arg	
	530					535					540					
Val	Ser	Gly	Val	Leu	Ser	Asn	Ala	Pro	Tyr	Leu	Leu	Asn	Val	Asp	Cys	
545					550					555					560	
Asp	His	Tyr	Ile	Asn	Asn	Ser	Lys	Ala	Leu	Arg	Glu	Ala	Met	Cys	Phe	

565								570					575			
Met	Met	Asp	Pro	Gln	Ser	Gly	Lys	Lys	Ile	Cys	Tyr	Val	Gln	Phe	Pro	
			580					585					590			
Gln	Arg	Phe	Asp	Gly	Ile	Asp	Arg	His	Asp	Arg	Tyr	Ser	Asn	Arg	Asn	
		595					600					605				
Val	Val	Phe	Phe	Asp	Ile	Asn	Met	Lys	Gly	Leu	Asp	Gly	Leu	Gln	Gly	
		610				615					620					
Pro	Ile	Tyr	Val	Thr	Gly	Cys	Val	Phe	Arg	Arg	Gln	Ala	Leu	Tyr	Gly	
		625			630				635							
Phe	Asp	Ala	Pro	Lys	Lys	Lys	Lys	Gly	Pro	Arg	Lys	Thr	Cys	Asn	Cys	
			645					650					655			
Trp	Pro	Lys	Trp	Cys	Leu	Leu	Cys	Phe	Gly	Ser	Arg	Lys	Asn	Arg	Lys	
		660					665					670				
Ala	Lys	Thr	Val	Ala	Ala	Asp	Lys	Lys	Lys	Lys	Asn	Arg	Glu	Ala	Ser	
		675				680					685					
Lys	Gln	Ile	His	Ala	Leu	Glu	Asn	Ile	Glu	Glu	Gly	Arg	Gly	His	Lys	
		690				695					700					
Val	Leu	Asn	Val	Glu	Gln	Ser	Thr	Glu	Ala	Met	Gln	Met	Lys	Leu	Gln	
		705			710				715				720			
Lys	Lys	Tyr	Gly	Gln	Ser	Pro	Val	Phe	Val	Ala	Ser	Ala	Arg	Leu	Glu	
		725					730					735				
Asn	Gly	Gly	Met	Ala	Arg	Asn	Ala	Ser	Pro	Ala	Cys	Leu	Leu	Lys	Glu	
			740				745					750				
Ala	Ile	Gln	Val	Ile	Ser	Arg	Gly	Tyr	Glu	Asp	Lys	Thr	Glu	Trp	Gly	
		755				760					765					
Lys	Glu	Ile	Gly	Trp	Ile	Tyr	Gly	Ser	Val	Thr	Glu	Asp	Ile	Leu	Thr	
		770				775					780					
Gly	Ser	Lys	Met	His	Ser	His	Gly	Trp	Arg	His	Val	Tyr	Cys	Thr	Pro	
		785			790				795				800			
Lys	Leu	Ala	Ala	Phe	Lys	Gly	Ser	Ala	Pro	Ile	Asn	Leu	Ser	Asp	Arg	
			805					810					815			
Leu	His	Gln	Val	Leu	Arg	Trp	Ala	Leu	Gly	Ser	Val	Glu	Ile	Phe	Leu	
		820				825					830					
Ser	Arg	His	Cys	Pro	Ile	Trp	Tyr	Gly	Tyr	Gly	Gly	Gly	Leu	Lys	Trp	
		835				840					845					
Leu	Glu	Arg	Leu	Ser	Tyr	Ile	Asn	Ser	Val	Val	Tyr	Pro	Trp	Thr	Ser	
		850				855					860					
Leu	Pro	Leu	Ile	Val	Tyr	Cys	Ser	Leu	Pro	Ala	Ile	Cys	Leu	Leu	Thr	
		865			870				875				880			
Gly	Lys	Phe	Ile	Val	Pro	Glu	Ile	Ser	Asn	Tyr	Ala	Ser	Ile	Leu	Phe	
			885					890					895			

Met Ala Leu Phe Ser Ser Ile Ala Ile Thr Gly Ile Leu Glu Met Gln
 900 905 910
 Trp Gly Lys Val Gly Ile Asp Asp Trp Trp Arg Asn Glu Gln Phe Trp
 915 920 925
 Val Ile Gly Gly Val Ser Ala His Leu Phe Ala Leu Phe Gln Gly Leu
 930 935 940
 Leu Lys Val Leu Ala Gly Val Asp Thr Asn Phe Thr Val Thr Ser Lys
 945 950 955 960
 Ala Ala Asp Asp Gly Glu Phe Ser Asp Leu Tyr Leu Phe Lys Trp Thr
 965 970 975
 Ser Leu Leu Ile Pro Pro Met Thr Leu Leu Ile Ile Asn Val Ile Gly
 980 985 990
 Val Ile Val Gly Val Ser Asp Ala Ile Ser Asn Gly Tyr Asp Ser Trp
 995 1000 1005
 Gly Pro Leu Phe Gly Arg Leu Phe Phe Ala Leu Trp Val Ile Ile His
 1010 1015 1020
 Leu Tyr Pro Phe Leu Lys Gly Leu Leu Gly Lys Gln Asp Arg Met Pro
 1025 1030 1035 1040
 Thr Ile Ile Val Val Trp Ser Ile Leu Leu Ala Ser Ile Leu Thr Leu
 1045 1050 1055
 Leu Trp Val Arg Val Asn Pro Phe Val Ala Lys Gly Gly Pro Ile Leu
 1060 1065 1070
 Glu Ile Cys Gly Leu Asp Cys Leu
 1075 1080

<210> 6
 <211> 56
 <212> PRT
 <213> Populus tremuloides

<400> 6

Met Met Glu Ser Gly Ala Pro Ile Cys His Thr Cys Gly Glu Gln Val
 1 5 10 15
 Gly His Asp Ala Asn Gly Glu Leu Phe Val Ala Cys His Glu Cys Ser
 20 25 30
 Tyr Pro Met Cys Lys Ser Cys Phe Glu Phe Glu Ile Lys Glu Gly Arg
 35 40 45
 Lys Val Cys Leu Arg Cys Gly Ser
 50 55

<210> 7
 <211> 5
 <212> PRT

<213> Unknown

<220>

<223> consensus sequence

<400> 7

Gln Val Leu Arg Trp

1

5